

Perceptual Modeling of Media Signals Based on Local Contrast and Directional Edges

Related Application Data

This patent application is a continuation of U.S. Patent Application No. 09/737,609 (allowed), filed December 13, 2000 (published as US 2002-0015508 A1). The 09/737,609 application is a continuation-in-part of U.S. Patent Application No. 09/596,658 (now U.S. Patent No. 6,631,198), entitled Perceptual Modeling of Media Signals Based on Local Contrast and Directional Edges, filed on June 19, 2000, by Hannigan, Reed, and Bradley. Each of the above patent documents is herein incorporated by reference.

The subject matter of the present application is related to that disclosed in US Patent 5,862,260, and in co-pending applications 09/503,881, filed February 14, 2000; which are hereby incorporated by reference.

Technical Field

The invention relates to multimedia signal processing, and in particular relates to perceptual modeling of media signals, such as images, video and audio.

Background and Summary

Perceptual modeling is often used in media signal processing applications to assess the extent to which changes to a media signal are noticeable to the human eye or ear. A perceptual model analyzes a media signal's ability to hide or mask changes. Such models are used in lossy signal compression and digital watermarking to minimize the perceptibility of these processes to a viewer or listener of the processed signal.

Lossy signal compression typically quantizes signal samples to reduce the memory or bandwidth required to store and transmit image and audio signals. Media signal codecs, like those defined in MPEG standards, use perceptual models to

identify parts of media signal that can be more heavily compressed while staying within a desired quality.

Digital watermarking is a process for modifying media content to embed a machine-readable code into the data content. The data may be modified such that the embedded code is imperceptible or nearly imperceptible to the user, yet may be detected through an automated detection process. Most commonly, digital watermarking is applied to media such as images, audio signals, and video signals. However, it may also be applied to other types of data, including documents (e.g., through line, word or character shifting), software, multi-dimensional graphics models, and surface textures of objects.

Digital watermarking systems have two primary components: an embedding component that embeds the watermark in the media content, and a reading component that detects and reads the embedded watermark. The embedding component embeds a watermark signal by altering data samples of the media content. The reading component analyzes content to detect whether a watermark is present. In applications where the watermark encodes information, the reader extracts this information from the detected watermark.

In digital watermarking, one aim is to insert the maximum possible watermark signal without significantly affecting signal quality. Perceptual models may be used to determine how to embed the watermark in a host media signal such that signal masks the watermark. In image watermarking, a watermark embedder can take advantage of the masking effect of the eye to increase the signal strength of a watermark in busy or high contrast image areas. However if this is done for all high frequency areas, a visually objectionable watermark or 'ringing' may become visible on connected directional edges.

In audio watermarking, busy or high contrast segments of an audio signal tend to have a greater masking effect. However, embedding a watermark in portions of an audio signal that represent pure tones may make the watermark more audible.

The invention provides methods for perceptual analysis of media signals. While particularly adapted to image signals, the invention applies to other types of media signals as well. One aspect of the invention is a method for perceptually analyzing a media signal to reduce perceptible artifacts around directional edges. The method
5 analyzes the media signal to compute a measure of directional edges. Based at least in part on the measure of directional edges, the method computes control data used to control changes to the media signal in a manner that controls perceptibility of the changes around directional edges.

For digital watermark applications, this method may be used to reduce perceptible
10 artifacts around connected edges. The method may also be used to reduce artifacts around directional edges in lossy signal compression schemes.

Another aspect of the invention is a method for perceptual analysis of a media signal based on local contrast. This method analyzes the media signal to compute measures of local contrast at samples within the media signal. Based at least in part
15 on the measures of local contrast, it computes a measure of visual sensitivity to changes of the media signal at the samples. To compute visual sensitivity to the local contrast, it applies a human visual model that relates local contrast to visual sensitivity.

In one implementation, the human visual model performs a non-linear mapping
20 function that is tuned to the eye's sensitivity to local contrast. In a plot of visual sensitivity versus contrast, visual sensitivity initially increases with contrast and then decreases. The mapping function exploits this attribute.

As in the case of the perceptual analysis based on directional edges, the perceptual analysis based on local contrast may be applied to a variety of media signal
25 processing applications. Some examples include digital watermarking and lossy signal compression.

The perceptual analyses based on directional edges and local contrast may be used independently, in combination with each other, or in combination with other perceptual models.

Further features will become apparent with reference to the following detailed description and accompanying drawings.

Brief Description of the Drawings

5 Fig. 1 is a diagram illustrating a method for perceptual modeling of a media signal.

Fig. 2 is a diagram illustrating a perceptual model for image signals.

Fig. 3 is a plot depicting a non-linear mapping function used to map local contrast to human sensitivity in a human perceptual model.

10

Detailed Description

Fig. 1 is a diagram illustrating a method for perceptual analysis of a media signal. The following discussion describes this method as applied to still image signals. However, the principals of the method are applicable to video and audio signals as well. This method includes perceptual modeling based on local contrast and directional edges. The result of this perceptual modeling may be combined with other forms of perceptual modeling. In addition, perceptual modeling based on local contrast and directional edges may be used independently.

15

The input to the perceptual analysis is a media signal 100, such as an image or audio signal. For the sake of an example, we describe an implementation for still images. In this case, the media signal is an image or part of an image. One aspect of the perceptual analysis 102 computes a measure of directional edges (104) at positions throughout the media signal. The method uses this measure to compute data to control changes to the input signal in a manner that reduces perceptibility of those changes. For example, the control data may be used to suppress a change to a sample or set of samples of a media signal as a function of the measure of directional edges at the position of the sample or samples in the media signal.

20

25

Another aspect of the perceptual analysis 106 computes a measure of local contrast at positions in the media signal. It then computes perceptual sensitivity at these positions based on the local contrast measurements and a perceptual model that models human sensitivity to contrast.

5 The perceptual analysis uses the results of the directional edge and local contrast perceptual modeling to compute a control vector (110). Elements of the control vector correspond to samples of the media signal. The magnitude of these elements reflect the relative impact that changes to corresponding samples are expected to have on perceptibility. A larger element value in the control vector means that changes to
10 a media signal at the position of that element are less likely to be noticeable, and thus, can tolerate greater changes for a desired amount of perceptibility. A smaller element value, conversely, means that changes will have a greater impact on perceptibility.

15 The perceptual analysis may combine either or both of the local contrast and directional edge measurements with other perceptual analyses data to compute the control vector (110). In the example of an image signal, the perceptual analysis may also compute a measure of image activity. Parts of an image that are highly busy or textured can withstand more changes for a desired amount of perceptibility relative to less busy, smoothly varying parts.

20 One way to perceptually analyze such signal activity is to high pass filter parts of the signal to measure the high frequency content of each part. The amount of high frequency components in a given part of the signal means that the part is more busy and likely to withstand more changes for a desired amount of perceptibility.

25 Another way to analyze signal activity is to measure the edges or sharp transitions per unit of the signal. A high measure of edges over a given area tends to indicate greater signal activity, and thus, a greater tolerance for changes for a desired amount of perceptibility. The exception, as noted above, is that directional edges are more sensitive to changes. Thus, a general measure of edginess without concern for directional edges will roughly indicate the extent to which a signal is perceptually insensitive to changes. A measure of directed edges over the same part of the signal

indicates the extent to which the signal has directional edges that are sensitive to changes. In a similar manner, the watermark detector should ignore areas with directional edges thus reducing the jamming effect of text and other strong directional edges.

5 In computing the control vector, the perceptual analysis takes into account the local contrast measure, the directional edge measure, and possibly other perceptual modeling such as models based on signal activity. Each of these perceptual analyses contributes to a composite control vector. Depending on the application, the perceptual analysis process may apply additional post processing to the composite
10 vector to generate a final control vector 112. This final control vector 114, or intermediate control vectors from the local contrast or directional edge analyses, may then be used in a variety of applications of perceptual modeling.

 One such application is the embedding of a digital watermark. For example, a control vector may be used to control the strength with which a watermark signal is
15 embedded in the media signal. The control vector can be used to adapt the watermark to the host signal in which it is embedded. This perceptual analysis method applies broadly to watermark methods that embed a watermark by changing the host signal in a temporal or spatial domain in which the signal is perceived (viewed or heard) or by changing the host signal in a transform domain, such as modifying transform
20 coefficients, subband samples, etc.

 For example, some watermark methods transform a host signal to a transform domain, modify transform coefficients or samples, and then inverse transform the modified coefficients or samples to produce a watermarked signal. Some examples include methods that modify Discrete Cosine Transform, Discrete Wavelet
25 Transform, or Discrete Fourier Transform coefficients. Elements of the control vector may correspond to parts of the host signal that are transformed to the selected transform domain. For example, in watermark process that encodes auxiliary information in DCT blocks, the elements of the control vector may correspond to the strength of watermark encoding in the DCT blocks. In a watermark process that

encodes auxiliary information in subband samples, the elements of the control vector may correspond to subband samples or groups of subband samples.

Another such application is lossy data compression of media signals. For example, a control vector may be used to control quantization of media signal samples in lossy compression schemes for images (e.g., JPEG, JPEG 2000), video (MPEG, H263, Windows Media Video), and audio (MPEG, AAC, Qdesign, Windows Media Audio, Twin VQ, ATRAC3, Dolby Digital AC-3, ePAC). As noted above, elements of the control vector may correspond to samples of the media signal, or transform coefficients or samples.

The granularity of the control vector may vary with the application and media signal type. For image signals, elements of the control vector may correspond to a pixel or blocks of pixels at a given spatial resolution. For audio signals, the elements of the control vector may correspond to an audio sample or frame of temporally contiguous audio samples.

The control vector may also correspond to media signal samples in a transform domain. An audio signal may be transformed into a time-frequency domain and then analyzed using aspects of the perceptual model described above. For example, an analysis tool may transform overlapping, temporal frames of an audio signal into a time frequency space, where the time axis corresponds to temporal frames, and the frequency corresponds to frequency coefficients for each frame.

The control vector may be used as a gain vector with elements that are used to adjust the strength of corresponding signal samples. For example, the elements may be used to adjust the signal strength of corresponding samples or groups of samples of a watermark signal.

Fig. 2 is a block diagram illustrating an implementation of a perceptual analysis for image signals. The inputs to the perceptual analysis include images or image blocks. In particular, the image input includes two versions of the same image at two different resolutions 200, 202. The resolution of the image may be obtained from the header file or may be estimated from the image itself. The format of the image at this

stage depends on the application. In this particular example, the perceptual analysis operates on luminance samples. The luminance samples may be generated by mapping color vector samples in an image from color space representations like RGB or CMYK to luminance values. The desired resolution of the image may be obtained
5 by up or down-sampling the image.

An initialization process 204 sets up blocks of the image at two different resolutions. In this case, one resolution (resolution x) is double the other (resolution y). The model applies the higher resolution block to directional edge mask 206 and edge strength detector 208. The directional edge mask measures directional edges in
10 a local neighborhood around each image sample. In particular, it computes the edge in several directions around a center sample in the neighborhood. The edge strength is calculated in four directions (horizontal, vertical, and along two diagonals), using the appropriate Gabor filters. The pixel is marked as a directional edge if the edge in one direction is significantly higher than the average of the other edge directions.

15 The edge strength detector 208 measures the strength of edges over the same neighborhood of image samples. One way to implement the edge strength detector is to apply a Laplacian filter to each neighborhood. The filter computes the dot product of the samples in the neighborhood with a two-dimensional array of scale factors (e.g., in a three by three window of samples, the center element has a peak value
20 surrounded by elements of a constant, negative value such as

-1 -1 -1

-1 8 -1

-1 -1 -1

Pratt, 'Digital Image Processing', p482, 1978).

25 Next, the model combines 210 corresponding elements of the edge mask and strength of edge calculations. In particular, it multiplies corresponding elements together. It then smooths the result by down sampling the resulting vector (e.g., down sample by 2) (212).

The model then applies a filter that grows directional edges 214. The effect of this filter is to retain directional edges and connect directional edges that are slightly disconnected. In effect, this process estimates the extent to which the directional edges are connected. One way to accomplish this effect is to apply an order filter
5 over a neighborhood of samples and choose an element less than halfway deep in the ordering from large to small values (e.g., five by five window choosing element 10). At this stage, the perceptual analysis of directional edges has generated control data, and particularly, a control vector representing a measure of directional edges. This vector may then be applied to selectively suppress the strength of a watermark signal
10 where directional edges are stronger.

Another aspect of the perceptual analysis measures local contrast, and maps the local contrast to a control data representing visual sensitivity. A local contrast analyzer 216, in this example, operates on the lower resolution version of the input image. It measures the local contrast in a neighborhood around each image sample in
15 that image. There are many different types of filters that may be used to measure local contrast. One such example is to compute the absolute value of the difference between the center element and each of eight surrounding elements, and then average the differences.

Next, the perceptual analysis maps the local contrast measurements to control
20 values based on a perceptual model 218 that simulates the eye's sensitivity to contrast. Fig. 3 illustrates a plot showing an example of the perceptual model. The perceptual model is depicted as a mapping function that maps local contrast values to a corresponding sensitivity values. These sensitivity values may act as control data, or may be converted to control data, used to adjust changes to the image.

25 For example, the control data for the image may comprises a control vector with elements that represent sensitivity: larger values mean low sensitivity, while smaller values mean high sensitivity. The mapping function follows the human eye's sensitivity to contrast. The vertical axis corresponds to a gain boost, meaning that larger values reflect that the image can tolerate more changes for a desired level of

perceptibility. The horizontal axis is a log scale of contrast. From Fig. 3, one can see that the eye is more sensitive to small levels of contrast, than no contrast. As the contrast increases, however, the eye becomes increasingly less sensitive to changes.

The increase in signal strength in the presence of a reference signal before being visually perceptible is a non-linear function (Barten, 'Contrast Sensitivity of the Human Eye', p.139, 1999). For watermarking applications, the mapping function has been derived experimentally, by applying a watermark signal at different strengths on top of a textured image of different contrasts. The strength at which the mark was just noticeable at each contrast was then determined visually, to generate a contrast versus watermark gain control curve.

The result of remapping the local contrast measurements is a control vector that adjusts changes to an image based on sensitivity. Elements of the vector corresponding to less sensitive image areas increase the change, while elements of the vector corresponding to more sensitive areas decrease the change.

The perceptual analysis combines the control vector from the contrast and directional edge modeling to form a gain vector. Additional post processing may then be applied to the gain vector. In the example shown in Fig. 2, the model applies a filter that removes isolated directional edges. To accomplish this, the perceptual analysis depicted in Fig. 2 uses a filter that suppresses random spikes in the gain vector. One such filter is a generalized order filter. One implementation of this type of filter orders elements from large to small values in a window around each element of the gain vector and replaces the center element with an element near the top (e.g., in an ordering of elements in a five by five window, choosing element 4).

For optimal performance for a particular applications, the model can be tuned by selecting combinations of filters that compliment each other and fine tuning the parameters of each filter.

The components of the perceptual analysis shown in Fig. 2 may vary from one implementation to another. Experiments indicate that a similar implementation to the one shown in Fig. 2, without the direction edge filter, may provide better results.

In some implementations, the re-mapping function applied to local contrast may not be a log mapping, but instead, some other non-linear mapping. Processes such as the just noticeable difference tests described above may be used to experimentally derive the non-linear mapping function for a particular application and type of media signal.

5 The perceptual analysis performs a contrast measurement and a directional edge measurement. In one implementation, it combines the two measurements and re-maps the result to take into account the human's perception of contrast. In an alternative implementation, the re-mapping may be applied to the contrast measurement, without a directional edge measurement.

10 ***Modified Implementation***

This section describes a modified implementation of the method shown in Fig. 2. This method for determining the magnitude of watermark signal to be applied based on local image characteristics can be described in four conceptual stages: directional edge detection, local contrast measurement and correction, combination of edge
15 detection and contrast, and non-linear contrast to gain mapping.

In this implementation, host image sample data is first fed into two separate stages – the contrast measurement stage and the directional edge detector stage. The results from these two stages are then combined to form a "corrected contrast measurement map" which reports the local contrast values for the image while protecting
20 directional edges. The results from this stage are then passed into a non-linear contrast to gain mapping stage which calculates the gain values, or magnitude of watermark signal varied image regions should receive.

Contrast Measurement

25 If a watermark is applied at equal strength throughout an image it will tend to be more visible in texturally flat regions, and less visible in busier areas. Conversely, it is usually more difficult to extract the watermark from pixels in busy regions than in those that are texturally flat. For these two reasons it is in general desirable to measure the textural contrast of the image to be watermarked on a local basis. The

obtained measurement is then used to control the strength of the watermark applied to the measured region. The process would be repeated for all regions of the image that require a watermark.

5 The measurement of contrast typically involves one or more filtering operations, possibly non-linear. To make an initial measurement of contrast, this implementation uses a band pass filtering operation. Although straightforward filtering produces a good initial result, refinement is used before the textural contrast measurement can be mapped to watermark strength without undue ill effects.

10 Using filtering to determine textural contrast, and hence watermark strength, without taking into account certain natural image characteristics leads to apocryphal indicators of how heavily a watermark should be applied to a given region. A region that contains a sudden transition in luminance may be labeled as a prime candidate for high watermark strength after initial contrast filtering. Regions that contain borders, text, or fine lines are some examples. If the region is heavy-handedly doused with a watermark it may appear objectionable depending upon the characteristics of the region and others that surround it. We classify as false contrast regions those that cannot truly support high watermark strength when the filter-based contrast measurement would indicate otherwise.

15 One method of dealing with potentially false contrast regions is to de-emphasize, or even penalize, such regions if they have an uncharacteristically high contrast. For example, we have been able to characterize our contrast filter by applying it to images we would regard as busy; the image does not degrade noticeably under high watermark strength. We found that on average the contrast measurement is relatively low compared with many of the false contrast regions. By characterizing our filter we set a peak expected contrast. If a region's contrast supersedes the expected peak contrast, its final assigned contrast is clipped at the expected peak value, or in some case reduced below the peak value. Although beneficial, the described contrast adjustment procedure works only on a local image region basis. By taking into account groups of regions, more intelligent decisions can be made regarding the

application of watermark strength. Our so-called directional edge finding method serves to do just that.

Connected Edge Detection

Edge detection algorithms have been well studied and evaluated in image processing literature. See J. S. Lim, Two-Dimensional Signal and Image Processing, pp. 476-495, PTR Prentice Hall, New Jersey, 1990; and W. K. Pratt, Digital Image Processing, pp. 478-492, John Wiley & Sons, New York, 1978.. Typical edge detection processes define an edge as a "step discontinuity in the image signal", and attempt to locate edges by convolving the image with a kernel that approximates a first or second derivative. See, for example, P. Kovesi, "Lecture 6, Classical Feature Detection,"

<http://www.cs.uwa.edu.au/undergraduate/courses/233.412/Lectures/Lecture6/lecture6.html>, The University of Western Australia, 2000.

Using a first derivative kernel, an edge occurs at local maxima, while for second derivatives, edges occur at the zero crossings. John F. Canny developed a standard in edge detection. See, J.F. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Analysis and Machine Intelligence, 8, pp. 679-698, 1986. This edge detection method starts by convolving an image with a 2-D Gaussian filter, and then differentiating this smoothed image in two orthogonal directions. By calculating the derivative in two orthogonal directions, one can determine the overall gradient direction and amplitude. Using this knowledge, the implementation then suppresses points which were non-maxima and values that were not local peaks. The final step involves thresholding the edges. Two threshold values are used. The first threshold, which is larger than the second, identifies most of the true edges. However, some discontinuities in edges may occur using this higher value, so pixels which are connected to these high threshold edges are also considered edges if they are above the second, smaller threshold. The resulting map typically provides a very good representation of the edges in the image.

The edge map provided by the Canny edge detection algorithm does not attempt to differentiate between connected edges and areas with a high concentration of random edge like textures. For example, for an images with random edge like textures like images of hair or fur, a Canny edge algorithm might construct the edge map in which the image portions depicting hair or fur are said to contain a high concentration of edges. However, the purpose of the edge map in our approach is to highlight connected edges that we should avoid when increasing watermark signal gain. The fur area should be able to hold a good deal of watermark signal, since the so-called edges are really a somewhat random texture, and placing a noise like signal in this texture is unlikely to be noticed. Thus, we modify Canny's edge detection algorithm to ignore random, closely packed edges in the following fashion. First, we take the edge map provided by Canny's algorithm and smear it with a 5x5 or 7x7 low pass filter kernel. This causes closely packed edges to bleed into one another. Next, we thin this smeared edge map using a min filter of a slightly higher order than the smear. This causes edges which were stretched to contract back in tighter than the original Canny edge mask. A composite, binary edge mask is then constructed by saying a pixel is on an edge if and only if the original Canny edge mask says it is an edge, but the min-filtered, smeared Canny Edge mask says it is not. This operation essentially allows edges that are boundary edges to remain, while closely packed edges disappear. The final step grows the edge map by a suitable radius to protect all pixels which may inadvertently be called "high contrast" areas due to these edges.

Combination of Edge Detection and Contrast Measurement

The third stage of our algorithm combines the results of the edge detection and contrast measurement stages. As stated above, the general method in determining the amount of watermark gain a certain image area should receive is mapping local contrast to some gain value. The purpose of the edge detection section is to identify areas which may be perceived as high contrast areas by the contrast algorithm, but in reality can not hold the watermark gain generally associated with such a high

contrast. Examples of these regions are text edges, object boundaries, and other directional edges.

This combination stage therefore takes as input the contrast map of the image as well as the binary edge map. For any contrast value which is not said to be on an edge, the contrast value is untouched. For those contrast values which are found to be on a directional edge however, the reported contrast value is calculated as a percentage of the original. We have found experimentally that a contrast reduction of 50% to 80% on edges provides a clean gain map with edges free of objectionable watermark ringing.

Non-Linear Contrast to Gain Mapping

The combined detail gain contrast calculation is then passed through a one - dimensional mapping function to obtain a final detail gain.

The detail gain function has a dip at low contrast, and is then approximately linear on a log scale. The function was calibrated for our application by using a mid-gray patch with white noise of various contrast levels, and embedding at different strengths until the watermark is just perceptually visible in order to build up the curve shape. Further tests were then performed on a standard image set. The shape obtained was very similar to the generalized contrast discrimination model reported in Peter G.J. Bartens, 'Contrast Sensitivity of the Human Eye and Its Effects on Image Quality', p.147-151, SPIE Press, 1999.

The model shows that the peak sensitivity of the human visual system to a signal, in the presence of a reference signal, is at a low contrast of about 0.9 % modulation of the reference signal.

The resulting detailed gain values are used to control the strength of the watermark signal embedded in image samples or groups of samples at corresponding locations in an image signal. In particular, the gain values are used to scale the amplitude or energy of the watermark signal at corresponding image sample locations in a host image. For more information, see Hannigan, B. Reed, A. and Bradley, B.,

Digital Watermarking Using Improved Human Visual System Model, which is attached as Appendix A and incorporated by reference.

Concluding Remarks

Having described and illustrated the principles of the technology with reference to specific implementations, it will be recognized that the technology can be implemented in many other, different, forms. To provide a comprehensive disclosure without unduly lengthening the specification, applicants incorporate by reference the patents and patent applications referenced above. Processes and components described in these applications may be used in various combinations with processes and components described above.

The methods and processes described above may be implemented in hardware, software or a combination of hardware and software. For example, the process may be incorporated into a watermark or media signal encoding system implemented in a computer or computer network. The methods and processes described above may be implemented in programs executed from the system's memory (a computer readable medium, such as an electronic, optical or magnetic storage device.)

The particular combinations of elements and features in the above-detailed embodiments are exemplary only; the interchanging and substitution of these teachings with other teachings in this and the incorporated-by-reference patents/applications are also contemplated.

Appendix A

Digital watermarking using improved human visual system model

Brett T. Hannigan, Alastair Reed, Brett Bradley

Digimarc Corporation, 19801 SW 72nd Ave., Suite 250, Tualatin, OR 97062

ABSTRACT

In digital watermarking, one aim is to insert the maximum possible watermark signal without significantly affecting image quality. Advantage can be taken of the masking effect of the eye to increase the signal strength in busy or high contrast image areas. The application of such a human visual system model to watermarking has been proposed by several authors. However if a simple contrast measurement is used, an objectionable ringing effect may become visible on connected directional edges.

In this paper we describe a method which distinguishes between connected directional edges and high frequency textured areas, which have no preferred edge direction. The watermark gain on connected directional edges is suppressed, while the gain in high contrast textures is increased. Overall, such a procedure accommodates a more robust watermark for the same level of visual degradation because the watermark is attenuated where it is truly objectionable, and enhanced where it is not.

Furthermore, some authors propose that the magnitude of a signal which can be imperceptibly placed in the presence of a reference signal can be described by a non-linear mapping of magnitude to local contrast. In this paper we derive a mapping function experimentally by determining the point of just noticeable difference between a reference image and a reference image with watermark.

Keywords: Digital watermarking, Edge Detection, Connected Edges, Human Visual System, Digimarc

1. INTRODUCTION

5

The process of digital watermarking involves tweaking pixel values at various regions within an image to encode some piece of information. This encoded information can be used to provide copyright information, to prevent illegal duplication, or even as a dynamic link between the image and online digital data. For most applications, the image owner would like to make the encoded data robust enough to ensure its detection while maintaining the high quality of the original image. Often times these two desires are in direct conflict with each other – any increase in robustness comes at the expense of increased signal visibility. However, the visual impact a signal has on a particular image is not solely related to the magnitude of the signal, nor is the detectability of a signal solely related to its magnitude. Local image characteristics also help determine the visibility of a signal and its detectability. For instance, a signal embedded into an image area which is primarily flat will be more visible than a signal of equal magnitude embedded into an area of high contrast. This latter signal may be more difficult to detect however, as areas of higher contrast can interfere or jam a signal more easily than a flat area.

10

15

20

Many watermarking schemes attempt to analyze images in order to identify how much watermark signal various image areas can hold while maintaining equal visibility. A related methodology attempts to identify how much watermark signal various image areas require to maintain equal detectability across the entire image. These techniques often accomplish this by calculating a value of local contrast, and mapping increasing contrast values to increasing signal magnitudes[1],[2]. This can present a problem though, as directional edges separating two distinct objects in an image may be identified as high contrast areas. This results in the application of a higher strength watermark

25

signal around the connected edge, which causes an objectionable watermark ringing or halo.

In this paper, we introduce a method which better identifies areas of true high contrast texture while protecting connected directional edges. Additionally, we describe a more accurate mapping to be used when mapping local contrast to watermark gain. We then discuss results of utilizing the new algorithm, and compare those results to a previous method. We conclude by discussing possible extensions to this method for the future.

2. ALGORITHM DESCRIPTION

Our algorithm for determining the magnitude of watermark signal to be applied based on local image characteristics can be described in four conceptual stages, as shown in figure 1.

The image data is first fed into two separate stages – the contrast measurement stage and the directional edge detector stage. The results from these two stages are then combined to form a "corrected contrast measurement map" which reports the local contrast values for the image while protecting directional edges. The results from this stage are then passed into a non-linear contrast to gain mapping stage which calculates the gain values, or magnitude of watermark signal varied image regions should receive.

2.1 Contrast Measurement

If a watermark is applied at equal strength throughout an image it will tend to be more visible in texturally flat regions, and less visible in busier areas. Conversely, it is usually more difficult to extract the watermark from pixels in busy regions than in those that are texturally flat. For these two reasons it is in general desirable to measure the textural contrast of the image to be watermarked on a local basis. The obtained measurement is

then used to control the strength of the watermark applied to the measured region. The process would be repeated for all regions of the image that require a watermark.

The measurement of contrast typically involves one or more filtering operations, possibly non-linear. To make an initial measurement of contrast, our implementation uses a band pass filtering operation. Although straightforward filtering produces a good initial result, refinement is required before the textural contrast measurement can be mapped to watermark strength without undue ill effects.

Using filtering to determine textural contrast, and hence watermark strength, without taking into account certain natural image characteristics leads to apocryphal indicators of how heavily a watermark should be applied to a given region. A region that contains a sudden transition in luminance may be labeled as a prime candidate for high watermark strength after initial contrast filtering. Regions that contain borders, text, or fine lines are some examples. If the region is heavy-handedly doused with a watermark it may appear objectionable depending upon the characteristics of the region and others that surround it. We classify as false contrast regions those that cannot truly support high watermark strength when the filter-based contrast measurement would indicate otherwise.

One method of dealing with potentially false contrast regions is to de-emphasize, or even penalize, such regions if they have an uncharacteristically high contrast. For example, we have been able to characterize our contrast filter by applying it to images we would regard as busy; the image does not degrade noticeably under high watermark strength. We found that on average the contrast measurement is relatively low compared with many of the false contrast regions. By characterizing our filter we set a peak expected contrast. If a region's contrast supersedes the expected peak contrast, its final assigned contrast is clipped at the expected peak value, or in some case reduced below the peak value. Although beneficial, the described contrast adjustment procedure works only on a local image region basis. By taking into account groups of regions, more intelligent

decisions can be made regarding the application of watermark strength. Our so-called directional edge finding method serves to do just that.

2.2 Connected Edge Detection

5

Edge detection algorithms have been evaluated in image processing literature [3],[4].

Typical edge detection processes define an edge as a "step discontinuity in the image signal" [5], and attempt to locate edges by convolving the image with a kernel that

approximates a first or second derivative. Using a first derivative kernel, an edge occurs

10 at local maxima, while for second derivatives, edges occur at the zero crossings. John F.

Canny developed the current standard in edge detection [6]. His method starts by

convolving an image with a 2-D Gaussian filter, and then differentiating this smoothed

image in two orthogonal directions. By calculating the derivative in two orthogonal directions, one can determine the overall gradient direction and amplitude. Using this

15 knowledge, the algorithm then suppresses points which were non-maxima and values that

were not local peaks. The final step involves thresholding the edges. Two threshold

values are used. The first threshold, which is larger than the second, identifies most of

the true edges. However, some discontinuities in edges may occur using this higher

value, so pixels which are connected to these high threshold edges are also considered

20 edges if they are above the second, smaller threshold. The resulting map typically

provides a very good representation of the edges in the image.

The edge map provided by the Canny edge detection algorithm does not attempt to

differentiate between connected edges and areas with a high concentration of random

25 edge like textures. For example, for an image like "Mandrill", a Canny edge algorithm

might construct the edge map in **figure 2a**, in which the fur around the mandrill's nose is

said to contain a high concentration of edges. However, the purpose of the edge map in

our algorithm is to highlight connected edges we should avoid when increasing

watermark signal gain. The fur area should be able to hold a good deal of watermark

signal, since the so-called edges are really a somewhat random texture, and placing a noise like signal in this texture is unlikely to be noticed. Thus, we modify Canny's edge detection algorithm to ignore random, closely packed edges in the following fashion. First, we take the edge map provided by Canny's algorithm and smear it with a 5x5 or
5 7x7 low pass filter kernel. This causes closely packed edges to bleed into one another. Next, we thin this smeared edge map using a min filter of a slightly higher order than the smear. This causes edges which were stretched to contract back in tighter than the original Canny edge mask. A composite, binary edge mask is then constructed by saying a pixel is on an edge if and only if the original Canny edge mask says it is an edge, but
10 the min-filtered, smeared Canny Edge mask says it is not. This operation essentially allows edges that are boundary edges to remain, while closely packed edges disappear. The final step grows the edge map by a suitable radius to protect all pixels which may inadvertently be called "high contrast" areas due to these edges. The effect of this process on "Mandrill", as can be seen in **figure 2b**, is to keep the edges around the
15 mandrill's eyes, nose, and outer face, while disregarding the texture edges on his fur.

2.3 Combination of Edge Detection and Contrast Measurement

The third stage of our algorithm combines the results of the edge detection and contrast
20 measurement stages. As stated above, the general method in determining the amount of watermark gain a certain image area should receive is mapping local contrast to some gain value. The purpose of the edge detection section is to identify areas which may be perceived as high contrast areas by the contrast algorithm, but in reality can not hold the watermark gain generally associated with such a high contrast. Examples of these
25 regions are text edges, object boundaries, and other directional edges.

This combination stage therefore takes as input the contrast map of the image as well as the binary edge map. For any contrast value which is not said to be on an edge, the contrast value is untouched. For those contrast values which are found to be on a

directional edge however, the reported contrast value is calculated as a percentage of the original. We have found experimentally that a contrast reduction of 50% to 80% on edges provides a clean gain map with edges free of objectionable watermark ringing.

5

2.4 Non-Linear Contrast to Gain Mapping

The combined detail gain contrast calculation is then passed through a one-dimensional mapping function shown below in **figure 3** to obtain a final detail gain.

10 **Figure 3. Combined Contrast to Watermark Gain Mapping Function**

The detail gain mapping function has a dip at low contrast, and is then approximately linear on a log scale. The function was calibrated for our application by using a mid-gray patch with white noise of various contrast levels, and embedding at different strengths
15 until the watermark is just perceptually visible in order to build up the curve shape. Further tests were then performed on a standard image set. The shape obtained was very similar to the generalized contrast discrimination model reported in [7]. The model shows that the peak sensitivity of the human visual system to a signal, in the presence of a reference signal, is at a low contrast of about 0.9 % modulation of the reference signal.

20

3. RESULTS

3.1 Comparison of Two Watermark Gain Methods

25 In **figure 4a**, watermark gain was simply varied depending on local image contrast. For some connected edges, as in the road sign, this resulted in a visually objectionable watermark that appeared as a ringing on connected edges. The proposed watermark gain with edge suppression method, identifies these areas as connected edges, and suppresses the watermark signal in these regions. The result shown in **figure 4b** is an image with

much cleaner text edges, since the objectionable ringing on connected edges has been suppressed.

5 The non-linear curve that is used to apply the watermark gain based on contrast takes advantage of the masking effect of the human eye, allowing a higher strength signal to be put into high contrast areas. This could not be done in the past due to the objectionable ringing of the watermark on connected edges. Since the watermark gain is suppressed on connected edges where the visibility is most objectionable, the gain can be increased to a higher level in other busy areas that are able to support a high watermark signal without
10 visibility problems. Thus a reduction in overall watermark visibility is obtained at about the same detection rate.

The detection rate is not greatly affected by the watermark suppression on connected edges, since a very small fraction of the overall number of pixels in the image are
15 reduced in strength. Also, in many watermarking schemes edge regions are much more difficult to read in general. Thus the overall detection rate is not greatly affected by the watermark suppression on connected edges.

Figure 5a shows the standard mandrill image. **Figure 5b** shows the image after a
20 watermark has been applied using the new gain algorithm. A strong watermark signal has been applied without any significant visual problems being apparent around connected edges. Instead, most of the watermark is placed inside the mandrill's fur, an area of true high contrast.

25 3.2 Algorithm Limitations

While the algorithm developed in this paper succeeds in protecting most connected edges and false-high-contrast areas from receiving too much watermark gain, it does have some limitations. Due to the nature of the modified Canny edge detection algorithm, areas with

narrow, parallel lines will still be considered true-high-contrast areas. This will cause image areas such as the girl's hair in the image labeled **figure 6b** to receive a large watermark gain, which appears visually objectionable. Additionally, small, closely spaced text may still have areas which receive too much watermark gain, although not
5 nearly as much as in a simple contrast method. This is due to the fact that in the modified Canny edge detection algorithm, the smeared Canny edge map may smear text together, and the min filter will not shrink this area back inside the original text boundaries. This would leave some edges of the text with uncorrected high-contrast values.

10 4. EXTENTIONS AND CONCLUSION

The algorithm developed in this paper successfully uses a local contrast value to determine the watermark gain, while additionally suppressing gain at most false-high-contrast areas. In certain instances, the algorithm mistakenly allows a large watermark
15 gain in an area that is unable to imperceptibly hold it. Future versions of the algorithm should therefore work to better differentiate between random, high contrast texture, and more uniform, parallel line patterns, as well as closely spaced text.

ACKNOWLEDGEMENTS

20 The authors would like to thank Tony Rodriguez, Don Haaga, Lou Berkley, Eric Hudson, Steve Whaley, Matt Riley, Peter Harmon, Ben Bounketh, David Cushman, Brian MacIntosh, Steve Decker, Burt Perry, Geoff Rhoads and Joel Meyer of Digimarc for their help in contributions to this paper.

REFERENCES

1. C.I. Podilchuk, W. Zeng, "Image-adaptive watermarking using visual models," 1998.
- 5 2. J.F. Delaigle, C. De Vleeschouwer, B. Macq, "Watermarking algorithm based on a human visual model," *Signal Processing*. **66**, pp. 319-335, 1998.
3. J. S. Lim, *Two-Dimensional Signal and Image Processing*, pp. 476-495, PTR Prentice Hall, New Jersey, 1990.
4. W. K. Pratt, *Digital Image Processing*, pp. 478-492, John Wiley & Sons, New
10 York, 1978.
5. P. Kovesi, "Lecture 6, Classical Feature Detection,"
<http://www.cs.uwa.edu.au/undergraduate/courses/233.412/Lectures/Lecture6/lecture6.html>, The University of Western Australia, 2000.
6. J.F. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern
15 Analysis and Machine Intelligence*, **8**, pp. 679-698, 1986.
7. Peter G.J. Bartens, 'Contrast Sensitivity of the Human Eye and Its Effects on Image Quality', p.147-151, SPIE Press, 1999